国外期刊科学数据管理调查与分析

江洪[1] 刘敬仪[1,2]
[1] 中国科学院武汉文献情报中心 武汉 430071
[2] 中国科学院大学经济与管理学院图书情报与档案管理系 北京 100190

摘要：［目的/意义］为国内期刊开展科学数据管理提供思考与借鉴，并结合国内相关现状提出期刊科学数据管理框架以及合理的发展建议。［方法/过程］通过调研选取国外开展科学数据管理较好的15家期刊为研究对象，分析其在科学数据管理相关政策、科学数据提交、仓储平台和开放获取等四方面的特点。［结果/结论］得到如下启示：制定完整的科学数据管理政策；建立详细的科学数据提交流程；关联专业的科学数据仓储平台；积极促进科学数据的开放获取等建议。

关键词：期刊 科学数据 科学数据管理 框架
分类号：G255.2
DOI：10.13266/j.issn.0252-3116.2019.09.014

1 引言

随着科学研究进入第四范式——数据密集型研究，如何更有效地管理科学数据成为相关机构、人员等需亟待解决的问题，隶属于国际科学协会理事会（International Council for Science，简称 ICSU）的世界数据系统（World Data System，简称 WDS）在其《战略规划2014-2018：全球科学可信的数据服务》[1]中提到："当今世界的研究工作比以往更加国际化、跨学科化和数据化，这就需要严谨的数据管理、开放获取数据和有效的协作协调。与此同时，研究领域也在发生彻底的改变，政府和资助者的政策对研究人员提出了新的和迫切的期望，以便完全地、开放地、及时地共享数据。比如，研究人员必须识别可信的数据服务以便保存和传播自己在研究活动中生产的数据，从而根据需求准备数据管理计划。"国务院办公厅于2018年4月2日发布的《科学数据管理办法》中指出，为进一步加强和规范科学数据管理，在科学数据的采集、汇交与保存；共享与利用；保密与安全方面提出了明确要求[2]。而期刊作为收集、存储和传播文献的主要工具之一，面对科学环境的改变和利益相关者对数据提出的新要求，需意识到在目前科学研究过程中产生的科学数据的重要性，只有长期保存和有效传播科学数据，才可以更好地促进科学事业的向前发展。

2 国内外研究现状

2.1 国外研究现状

国外对于期刊科学数据管理的理论研究开展的较早，并且随着实践发展中相关模型的建立和越来越多的期刊对科学数据管理的探索，理论研究的探讨也越来越全面和具体，可主要分为以下三个方面：①相关制度研究：如 K. R. Wilhelmus 通过调研分析等方法对领域期刊开展科学数据管理提出应强制要求作者提交数据集的建议[3]；G. Santori 建议应积极推动数据再生性并签署相关声明[4]等；②已有理论模型研究：如 P. Sturges 对已形成的期刊科学数据管理相关模型如"期刊数据共享政策模型"[5]等进行展示与描述；③具体期刊研究：如 N. P. Gleditsch 等以"特殊数据特性"的形式发布数据集的《和平研究杂志》为实例进行阐述与分析[6]。在工作实践方面，国外许多知名期刊出版商如 Elsevier[7]和 Wiley[8]等下属的期刊皆对研究者提交的文章提出对应的科学数据管理要求，以促进科学数据的保存与利用；国际科学、技术和医学出版商协会 STM 也曾建立数据出版金字塔模型，为相关人员提供了使数据可被获取的通用方式，该金字塔模型共分为四层，从底层至顶层代表的含义依次是（其中②与③皆属于第二层）：①在起草人和机构硬盘中存储的数据；②从文章中引用的数据并保存在数据中心和仓储中；

作者简介：江洪（ORCID：0000-0003-3806-1856），副主任，研究员，E-mail：jianghong@mail.whlib.ac.cn；刘敬仪（ORCID：0000-0002-3462-6429），硕士研究生。
收稿日期：2018-09-29 修回日期：2018-11-21 本文起止页码：127-134 本文责任编辑：杜杏叶

③数据出版时描述数据集的可用性;④以任何补充文件的形式进一步解释数据;⑤文章中包含数据并加以解释。具体模型如图1所示[9]:

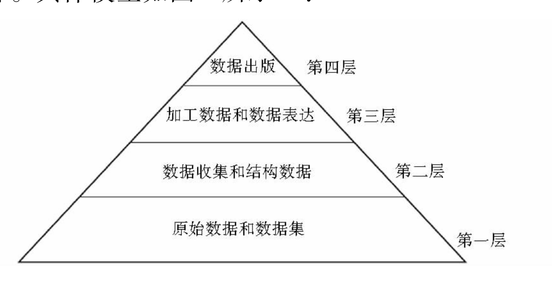


图1 STM数据出版金字塔模型

2.2 国内研究现状

在国内也有部分学者意识到期刊进行科学数据管理的重要性,但由于目前我国期刊科学数据管理发展尚处于萌芽阶段,因此在已有的理论研究中学者们多是选择国外较有代表性的期刊为研究对象,从不同的视角探究和分析国外期刊开展科学数据管理的特点,从而为国内期刊提供借鉴之处,主要可分为两个大类:①对支撑期刊开展科学数据管理的政策进行探讨:如黄如花和李楠通过调研国外5种纯数据期刊的政策提出我国应基于科研数据生命周期建立政策等[10];韩铭扬和姜鑫则对BioMed Central展开研究得出我国相关期刊应加强数据规范等建议[11];②针对期刊科学数据管理的某一环节如数据引用[12]或整体流程展开分析并指出我国期刊科学数据管理现存问题和需加强的具体方面等[13]。在工作实践方面,国内期刊在推动科学数据管理工作方面较落后于国外期刊,处于刚起步的阶段,《数据分析与知识发现》(原《现代图书情报技术》)于2016年12月要求作者提交支撑来稿论文结论的科学数据,并以适当方式实现数据的共享[14];同年同月,《图书馆杂志》为探索实现期刊论文数据的存储、管理等多个功能,建设了图书馆杂志数据管理平台(测试版),并于2017年正式进入试运行阶段[15]。

总体而言,国外期刊在开展科学数据管理相关工作方面早于国内期刊,发展相对较成熟,在此方面的理论研究也较多,值得国内相关研究者学习和借鉴。本文通过文献调研和扫描相关信息源的方法,选取了 G^3(Geochemistry, Geophysics, Geosystems 的简称)、*BioMed Central*、*Geology*、*Geosphere*、*GSA Bulletin*、*Lithosphere*、*Environmental and Engineering Geoscience*、*PLOS ONE*、*Earth System Science Data*、*Molecular Ecology*、PANGAEA期刊、*Nature*、*Geoscience Data Journal*、*Scientific Data*(该刊为纯数据期刊)、Elsevier期刊共15家国外期刊作为研究对象,从科学数据管理相关政策、提交要求、仓储平台和开放获取这四个方面探究并分析这些期刊开展科学数据管理的现状,以期为国内期刊开展相关工作提供思考。

3 国外期刊科学数据管理现状

3.1 科学数据管理相关政策

在开展科学数据管理工作时,相关政策的设置有利于期刊管理者、数据生产者、有关机构、资助方等利益相关者对科学数据进行有效管理,促进科学数据在其生命周期的每一个环节尽可能发挥其最大价值,从而推动科学事业的交流与发展。现将所调研期刊的科学数据管理相关政策整理成表,具体见表1(按可查到的相关政策发布时间由远至近排序)。

由表1可知,所有被调研的期刊都具有相关的科学数据管理政策,其中 G^3(1999年创刊[29])采用的由AGU制定的《AGU出版物数据政策》发布时间最早,在1993年便已被出版协会采用,可见政策制定方和采用方较早地意识到了期刊进行科学数据管理的重要性和必要性。在内容要素方面,所有政策皆包含数据保存、共享、获取、复用这四个方面的基本内容,具体如*Geoscience Data Journal*、*Scientific Data*等期刊都对数据引用的要求和原则展开了明确说明,并在期刊网页中设置"快速问答"栏目和编辑联系方式为作者答疑解惑,以此进一步完善和保证科学数据管理工作更有效地进行。此外,为服务于不同领域的作者,部分期刊政策的制定方联合期刊为作者提供最适合的相关政策,如施普林格·自然依照下设期刊的规模和网站风格将科研数据政策划分为4种类型[30],以期为不同领域的作者提供明确的指导,详细见表2。

3.2 科学数据提交要求

3.2.1 描述数据 为提高论文相关数据集的透明性和使用价值,期刊一般都会要求作者对所提交的数据进行完整的描述,并提供数据描述模板供用户下载使用以更好地起草和完成相关内容。调研中发现受较多期刊接受和采用的是Data Descriptor中对数据描述符做出的内容要素规范,其要求数据描述符原稿应包含以下9个方面的要素①标题:最大限度为110个字符,包含空格;②摘要:最大限度为170个词,无参考文献;③背景与概要:最大限度为700个词;④方法:不限长度;⑤数据记录:不限长度;⑥技术确认:不限长度;⑦使用注释:不限长度;⑧图:通常每篇原稿不多于3个;⑨表:通常每篇原稿不多于3个[31]。此外,也有期刊

江洪，刘敬仪. 国外期刊科学数据管理调查与分析[J]. 图书情报工作,2019,63(9):127－134.

表 1　期刊科学数据管理相关政策一览表

期刊名称	政策制定方	相关政策	内容要素	发布时间
*G*³	美国地球物理联合会(American Geophysical Union,简称 AGU)	《AGU 出版物数据政策》[16]	• 在经过检验的数据中心识别与保存数据 • 共享数据 • 描述数据或方法的详细信息 • 在文章致谢部分写明他人如何引用文中数据和获取限制 • 鼓励根据 FORCE11 和 ESIP 共用指导来引用数据集	1993 年 11 月被出版协会采用 2013 年被理事会采用 2016 年更新
BioMed Central	BioMed Central	《开放数据》[17]	• 依照《潘顿科学开放数据原则》 • 作者放弃版权和相关权利 • 遵循《知识共享 CCO 1.0 公开领域奉献豁免》	2013 年 9 月
Geology *Geosphere* *GSA Bulletin* *Lithosphere* *Environmental & Engineering Geoscience*	美国地理协会 The Geological Society of America, 简称 GSA)	《GSA 数据出版政策》[18]	• 地理科学数据保存 • 开放获取数据 • 推荐数据仓储	2014 年 2 月
PLOS ONE	*PLOS ONE*	《数据可用性》[19]	• 除极少情况,数据无任何限制被获取 • 提交 PLOS 的《数据可用性声明》 • 数据共享	2014 年 3 月
Earth System Science Data	FORCE11[20]	《数据引用原则联合声明》(*Joint Declaration of Data Citation Principles*, 简称 JDDCP)[21]	• 学术信用与规范和法律归属 • 证明 • 需唯一识别 • 获取 • 长久保存 • 具体和可验证性 • 互操作性与灵活性	2014 年
Molecular Ecology	Wiley	《数据共享政策》[22]	• 数据可获取性声明 • 公开存档 • 推荐合适仓储	未说明
	FORCE11	JDDCP	同 JDDCP 内容要素	2014 年
PANGAEA 期刊	由不同的利益相关者联合设计与认可[23]	《FAIR：科学数据管理的指导性原则》[24]	• 可检索性 • 可获取性 • 互操作性 • 可重用性	2016 年 3 月
Nature	Nature	《数据、材料、代码和协议可用性》[25]	• 数据可用性声明 • 数据引用政策 • 推荐数据仓储(figshare 和 Dryad)	于 2016 年 9 月发布
Geoscience Data Journal	*Geoscience Data Journal*	《作者指导》[26]	• 数据需在 GDJ 认可的仓储或数据中心正式存档(最好已分配指定的 DOI) • 使用 GDJ 在线提交系统 • 认可数据仓储列表 • 文章中的数据集部分以细节描述和 DOI 为特征 • 数据参考 • 数据引用	数据引用相关指导为 2017 年 10 月发布
Scientific Data	施普林格·自然[27]	《数据政策》[27]	• 数据存储政策 • 选择仓储 • 首次提交相关要求 • 出版要求 • 动物或人类主题实验数据 • 数据类型 • 数据保存 • 数据集更新 • 数据仓储标准 • 其他数据仓储推荐 • 数据引用(签署 JDDCP)	未说明

（续表1）

期刊名称	政策制定方	相关政策	内容要素	发布时间
Elsevier 期刊	Elsevier	《科研数据政策》[28]	● 共享科研数据 ● 获取与复用 ● 数据可用性说明 ● 开发工具和服务以支撑研究者发现、使用和复用数据 ● 为科研人员、机构和资助者提供有效管理数据所需的结构、工作流和技术 ● 继续参与行业项目、标准与政策机构	未说明

表2　施普林格·自然数据政策类型

数据政策摘要	示例期刊
鼓励数据共享与引用	《光合作用研究》
数据共享和对应凭证	《植物与土壤》
鼓励数据共享并需要数据可用性声明	《帕尔格雷夫——通信》
需要数据共享、数据共享凭证和数据的同行评审	《科学数据》

如 *Geoscience Data Journal* 要求作者在描述数据集的论文中给出其收集、处理和文件格式的细节等要素，并建议作者使用期刊提供的在线系统完成相关内容的编写和提交工作[26]。

3.2.2　可用性说明　调研中发现，大部分期刊会要求作者在提交数据的同时提交数据可用性说明，说明的描述要点包含数据的可获取性、检索号码、适用和限制条件等，从而确保原稿中的数据可以被其他研究者合理引用和获取，现列举部分期刊对数据可用性说明的要求如表3，其中需特殊说明的是 *Nature* 采用了针对数据可用性制定的指导政策——《数据可用性说明和数据引用政策：作者指导》以更好地帮助作者完成原稿中数据可用性说明的提交，该政策还附有详细示例辅助作者进一步理解政策内容。

表3　期刊对数据可用性说明要求（部分）

期刊名称	可用性说明内容描述要点	备注
Earth System Science Data	如何使得其他研究者获取附在原稿中的所有数据[20]	作者需将数据可用性说明放置在原稿的最后、"致谢"部分之前
Molecular Ecology	原稿中所有数据可用的检索号码或是 DOI[32]	若作者缺少数据可用性说明，则期刊编辑将有权拒绝原稿的提交请求
Nature	数据使用的适用条件、获取编码、其他特殊识别符、可公开获取数据集的相关网页链接和任何访问非公开获取数据集的条件[33]	当作者提供图源数据的时候，声明中也要涵盖描述要点中提到的内容

3.3　科学数据仓储平台

所调研的期刊皆鼓励作者在提交数据的同时将数据存入合适的数据仓储中，并且分配 DOI，使得数据更易被找到、复用和长期保存，从而进一步扩大影响力，如，GSA 鼓励其所有作者将出版论文中所使用的新数据在被信赖的仓储中进行永久保存，并列出五个方面的益处以促进作者完成该项工作：①保持对数据的开放获取；②提供长期保存；③使用持久稳定且独一无二的识别符；④可登记元数据；⑤具有质量保证[18]。现将调研期刊推荐较多的部分科学数据仓储平台整理成表，具体见表4。

表4所列的5个数据仓储平台横跨不同的学科领域，为提交数据的期刊作者提供数据保存、开放获取、分配 DOI、合理引用和发现等服务。其中，需说明的是 *Dryad* 指出当数据保存的老版本被遗弃时，将通过迁移通用文件格式的方法实现数据的长期保存，并与 *DataONE* 进行合作以保证使用者无限期获取数据内容。此外，*Mendeley Data* 提出会为数据提交者提供分享进度服务，鼓励研究者们将未出版的数据分享给同行或是资助机构以向前推进科研进程，并且提交者具有可完全控制有哪些人员可以访问和下载所提交数据的权利。

3.4　科学数据开放获取

期刊一般都通过自制开放获取相关原则与说明或采取被认可的相关原则与许可两种途径提倡、要求作者将提交的科学数据开放获取，以促进数据的流动、发现和引用。

3.4.1　自制开放获取相关原则与说明　AGU 在其2012年发布的立场声明中提到"地球与空间科学数据应该以多种格式被广泛获取，并且数据长期保存是科学家和赞助机构的责任"，为遵循这条声明并促进科学探索和发现，同时允许所有研究者对 AGU 上的所有结果进行全面评估，要求所有必要的数据必须是可被理解的、评估的、复制的，并且建立于报告研究基础上的（数据）必须是可用的和可获取的，因此鼓励作者在经检验的数据仓储或中心中识别并保存他们的数据，若

江洪，刘敬仪. 国外期刊科学数据管理调查与分析[J]. 图书情报工作,2019,63(9):127－134.

表4　科学数据仓储平台(部分)

名称	学科领域	主要特点
figshare	光学;生物学;化学;地理;社会科学等	• 使数据更易被发现并向所有读者开放; • 在所有类型浏览器中实现安全托管和可视化; • 无需担心文件大小或类型便可轻松上传文件; • 所有数据皆是可引用的并拥有 DOI; • 增加终端用户参与; • 管理并权衡你的数字文件影响力; • 满足资助方对数据的要求; • 托管和可视化数据[34]
Dryad	生态学;植物学;遗传学;进化等	• 灵活的数据格式; • 适合合作期刊的数据提交工作流,使数据提交变得容易; • 给期刊提供在同行评审期间私下获取数据的选项,允许提交者设置有期限的禁用后出版; • 数据与相应的出版物相关联,合适的情况下与专业的数据仓储相关联(例如 GenBank); • 给数据分配 DOI,从而研究人员可通过数据引用获取职业诚信; • 促进数据可视化; • 免费下载并无法律障碍地重用内容; • 管理数据以保证文件和元数据的准确性; • 提交者可更新数据文件; • 长期保存[35]
Mendeley Data	神经科学;生物学;生态学;地球与环境科学等	• 保存数据集; • 使科研数据可被他人引用; • 通过 CoreTrustSeal 认证; • 分享进度[36]
Protein Data Bank	分子生物学;结构生物学;计算生物学等	• 新的架构和服务使得获取信息更加快速; • 实施新举措以提升存档管理; • 已提升文本检索[37]
GenBank	分子生态学;生物医学等	• 开放获取; • 核苷酸序列和其蛋白质翻译的注释集合[38]

没有相关可用的公开仓储,并且这些数据是不容易被包含在一个附加补充当中的情况下,作者被期望在作品发表后的五年时间内管理上述数据,并提供一个透明的处理过程以便根据要求将数据提供给任何人,同时要求作者在论文的“致谢”部分有一个明确的说明以阐述其他用户如何从该文章获取数据(通过附加补充、仓储、其他来源等等)并说明在获取方面的任何限制,AGU 出版物不引用未被管理或不能被可靠地供任何人请求数据使用的数据集[16];*PLOS ONE* 则要求数据提交者必须同意共享数据、相关元数据和方法,若拒绝,则期刊有权不接受数据的提交请求,且需注意以下两种情况:①若作者未自己收集数据而是使用其他来源,那么该来源必须被合适地记录;②若在出版后,数据访问的限制被曝光,期刊保留发布修正的权利,以联系作者的机构和资助者,或在极端情况下撤回出版[19]。

3.4.2　采取被认可的相关原则与许可　BioMed Central 在数据开放获取方面采用了《潘顿科学开放数据原则》(*Panton Principles for Open Data in Science*),要求作者在出版数据方面放弃版权和相关权利从而使他人能够不需要特殊许可便能重新使用数据,并且在通常情况下(除非在个别的文章中另有说明),在该期刊上发表的开放获取文章中的数据都要遵循《知识共享 CC0 1.0 公开领域奉献豁免》(*Creative Commons CC0 1.0 Public Domain Dedication waiver*),任何人在重新使用该期刊上发表的数据时,尽管没有法律要求,但尽可能要在衍生作品中引用数据的来源[17];Molecular Ecology(同 Wiley 的要求)希望作者签署《开放获取协议》为出版作品提供一份知识共享许可(Wiley 采用的有:CC-BY, CC-BY-NC 和 CC-BY-NC-ND)从而符合开放获取出版条款并确保出版作品尽可能广泛传播[39]。

4　国内期刊科学数据管理建议

科学数据在当今科研环境中发挥着十分重要的作用,国际上也有越来越多的期刊要求作者在提交论文的同时提交与论文相关的科学数据,以推动科学研究更有效地交流。由此可见,数据出版是使数据可获取的重要方式,国内期刊也应跟上国际趋势,积极开展科学数据管理相关工作,推动数据出版。笔者根据调研对象在科研数据管理方面的特点绘制了期刊科学数据管理框架,如图 2 所示,并结合框架为国内期刊开展科学数据管理提出建议。

4.1　制定完整的科学数据管理政策

期刊在开展科学数据管理时,需要制定完整的科

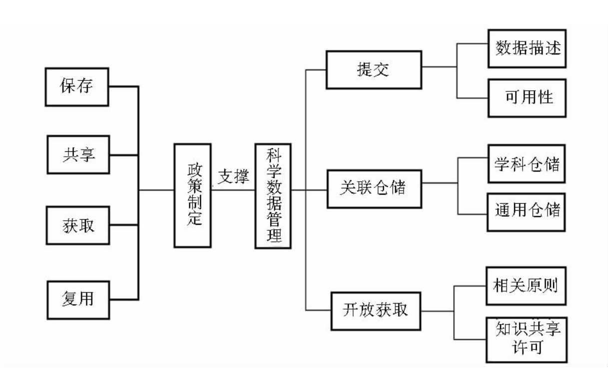


图 2 期刊科学数据管理框架

学数据管理政策支撑整个工作流程,使每个环节都有章可循,有政策可依,同时保护期刊、数据提交者和其他利益相关者的权益。建议完整的科学数据管理政策应包含:①保存:数据作为科学事业的重要产物,应被长期有效保存以备再次利用。数据提交者应选择合适的仓储(经期刊认可)存储数据;期刊有责任引导提交者保存数据并向其推荐有保证的、合适的数据仓储。②共享:开放科学环境,数据共享更能够促进数据流动和科学交流。数据提交者应意识到数据共享代表着自己的科学研究产物可分享给其他研究者,为他人制造福利,推动相关科学事业的发展;期刊应起到督促的作用,督促提交者将数据共享给他人,帮助推进科学研究进程。③获取:数据的可获取性保证数据是有效可用的,便于其他研究者验证与使用,提升数据价值。提交者应告知其他研究者通过何种方式(如数据仓储、相关链接等)可获取提交的数据,从而为其他研究者进行科学研究节省获取数据的时间;期刊在该环节应审查提交的数据是否可被有效获取。④复用:数据的复用需具备严格的引用规定以保证合理使用他人的科学产物。提交者应保证数据的可复用性,即提交的数据必须是真实的、有依据的;期刊应采用国际较通用的引用规定规范复用环节,如 JDDCP 等。以上四方面的内容即围绕数据生命周期制定完整的科学数据管理政策,政策一经制定,各利益相关者必须严格遵守,保证数据可在其每个环节都发挥最大价值。

4.2 建立详细的科学数据提交流程

科学数据应作为论文的一部分或是附属部分与论文一同提交,为此,期刊应建立详细的提交流程引导提交者进行有效提交,并且应注意避免因流程过于繁琐而影响该环节完成的质量,同时建议期刊配备专门人员或设置在线提交系统(带有合适的样例)帮助提交者完成此项工作。建议流程应重点包含:①对数据的清晰描述:清晰地描述数据便于其他研究者对数据的正确理解和选择,描述要素应涵盖标题、摘要、背景、方法、收集与处理、可信度确认、相关图表、文件格式等,同时期刊可根据自身特点对每个要素进行字符约束,避免数据描述过于冗长,这方面可参照 *Data Descriptor* 对数据描述提出的要求;②填写可用性说明:可用性说明的存在便于其他研究者获取并引用该数据,其应作为必须提交的一部分被放置于原稿的最后,包含其他研究者可获取该数据的检索码或 DOI、注明数据的适用条件和限制因素,从而保证数据更有效地被复用,具体可参照 PLOS ONE 与 Nature 采用的数据可用性的相关说明;③提交确认:期刊相关人员或在线提交系统需对提交的数据信息进行确认,防止疏漏影响后续共享、引用等工作。

4.3 关联专业的科学数据仓储平台

数据的长期保存对科学事业的发展十分必要,利于数据生产者存储大量的科学数据,并进行相关的管理工作。目前国际上已有很多建立较完善的数据仓储,通过先进技术实现对数据的长期保存,并可实现大容量保存数据,为数据生产者提供极大的便利。期刊若想更好地开展科学数据管理,应与合适的数据仓储进行合作,或将被期刊认可的数据仓储推荐给数据提交者,便于提交者进行合理选择,建议将数据提交者是否已将数据存储至数据仓储中作为该数据支撑的论文是否可在期刊上得以发表的必要条件之一。此外,期刊在关联科学数据仓储平台时需注意:①确定仓储平台可以很好地实现对数据的长期保存与安全存储,建议关联能通过及时迁移通用文件格式的方法实现数据长期保存的仓储,如 Dryad;②了解仓储平台接受的数据所属学科领域,建议期刊根据本身所涉及的学科领域关联对应的数据仓储,如地球与环境科学可使用 EarthChem[40],分子生物学方面推荐 Protein Data Bank,生物医学应使用 GenBank,学科涵盖较广泛的期刊推荐通用仓储如 figshare、Harvard Dataverse[41] 等;③清楚关联数据仓储的相关原则,期刊应比数据提交者先清楚仓储的相关要求和原则,从而更好地向提交者推荐,提高数据存储工作的效率。

4.4 积极促进科学数据的开放获取

当今的科研环境已走向开放,期刊应顺应大环境的变化积极促进科学数据的开放获取,以提高研究者获取信息的时效性,加快科学研究进程,同时可以在一定程度上扩大作者的影响力。为更好地促进开放获取,期刊需首要解决也是较难解决的便是知识产权和所有权的归属问题,就目前国际上较多期刊的做法来

看,建议期刊要求作者放弃版权和其他相关权利,以方便其他研究者毫无限制地复用数据,实现数据完全的开放获取,但作者可提出对要发布的数据设置一定期限内的访问限制(需作者列出具体原因,经编辑同意后限制生效),如可在数据发布的一年内禁止将数据公开,限制访问的人员等,再有例外情况,如涉及到人类主题的数据或者濒危物种位置的敏感信息等可由期刊编辑规定其他研究者获取该数据的条件或决定是否开放。此外,建议期刊自己制定或采用相关政策、原则对数据提交者在开放获取方面提出要求,可参照《潘顿科学开放数据原则》,同时需与提交者签署期刊认可的《开放获取协议》、知识产权许可(CC-BY, CC-BY-NC和CC-BY-NC-ND等)等,以保证数据开放获取的顺利施行,促进科学数据的广泛传播,推动科学事业的发展。

5 结语

科学数据作为科学研究的产物和推动科学事业发展不可缺少的组成部分,已成为当今科学界较热的关注点之一,如何对科学数据进行有效的管理,亦是诸多数据生产者、研究人员、基金机构等利益相关者一直探索的问题。期刊在过去相当长的一段时间对属于科学研究产物之一的论文进行了有效的保存和传播,因此更需意识到在越来越多的利益相关者对支撑论文的同时也是科学研究产物之一的科学数据提出需求的情况下,应积极开展科学数据管理,在这一方面,相比于国内期刊,国外期刊发展得较成熟,已具有较完整的科学数据管理体系,值得国内期刊借鉴与学习。下一步,建议国内期刊应主要从科学数据管理政策、数据提交流程、数据仓储平台和开放获取这四方面入手,全面开展科学数据管理,为促进国内乃至国际科学研究贡献力量。

参考文献:

[1] ICSU world data system. Strategic plans[EB/OL]. [2018－09－13]. http://www.icsu-wds.org/publications/strategic-plans/strategic-plan-2014-2018.
[2] 国务院办公厅. 国务院办公厅关于印发科学数据管理办法的通知[EB/OL]. [2018－09－14]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.
[3] WILHELMUS K R. The analysis of clinical research: mandatory submission of data sets journals should have access to research data[J]. Archives of ophthalmology, 2000,118(2):273－274.
[4] SANTORI G. Journals should drive data reproducibility[J]. Nature, 2016, 535(7612):355.
[5] STURGES P, BAMKIN M, ANDERS J H S, et al. Research data sharing: developing a stakeholder-driven model for journal policies[J]. Journal of the Association for Information Science & Technology, 2015, 66(12):2445－2455.
[6] GLEDITSCH N P, NORDÅS R, URDAL H. Peer review and replication data: best practice from journal of peace research[J]. College & research libraries, 2017, 78(3):267－271.
[7] Elsevier. Publishing with Elsevier: step-by-step[EB/OL]. [2018－09－18]. https://www.elsevier.com/authors/journal-authors/submit-your-paper.
[8] Wiley. Open data[EB/OL]. [2018－09－18]. https://authorservices.wiley.com/open-science/open-data/index.html.
[9] STM. Integration of data and publications[EB/OL]. [2018－09－21]. https://www.stm-assoc.org/standards-technology/integration-of-data-and-publications/.
[10] 黄如花, 李楠. 基于数据生命周期模型的国外数据期刊政策研究[J]. 图书与情报, 2017(3):36－42＋108.
[11] 韩铭扬,姜鑫. 国际开放存取期刊数据政策调查分析[J]. 情报杂志,2017,36(8):161－165.
[12] 迟玉琢, 王延飞. 面向科学数据管理的科学数据引用内容分析框架[J]. 情报学报, 2018,37(1):43－51.
[13] 刘灿,王玲,任胜利. 数据期刊的发展现状及趋势分析. 编辑学报,2018,30(4):344－349.
[14]《数据分析与知识发现》编辑部. 支撑数据提交要求[EB/OL]. [2018－09－17]. http://manu44.magtech.com.cn/Jwk_infotech_wk3/fileup/2096-3467/NEWS/20161213090914.pdf.
[15]《图书馆杂志》数据出版管理平台. 关于我们[EB/OL]. [2018－09－17]. http://data.libraryjournal.com.cn/home/static/about.jsp.
[16] AGU. AGU publications data policy[EB/OL]. [2018－09－18]. https://publications.agu.org/author-resource-center/publication-policies/data-policy/.
[17] BMC. Open data[EB/OL]. [2018－09－18]. https://www.biomedcentral.com/about/policies/open-data.
[18] The Geological Society of America. GSA data policy for publications[EB/OL]. [2018－09－18]. https://www.geosociety.org/gsa/pubs/dataPolicy.aspx.
[19] PLOS ONE. Data availability[EB/OL]. [2018－09－18]. https://journals.plos.org/plosone/s/data-availability.
[20] FORCE11. Joint declaration of data citation principles － final[EB/OL]. [2018－09－18]. https://www.force11.org/datacitationprinciples.
[21] Earth system science data. Data policy[EB/OL]. [2018－09－18]. https://www.earth-system-science-data.net/about/data_policy.html.
[22] Wiley. Wiley's data sharing policies[EB/OL]. [2018－09－18]. https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/data-sharing-policy.html.

[23] WILKINSON M D, DUMONTIER M, AALBERSBERG I J, et al. The FAIR guiding principles for scientific data management and stewardship[J]. Scientific data, 2016, 3:167 - 172.
[24] PANGAEA. About[EB/OL]. [2018 - 09 - 18]. https://www.pangaea.de/about/.
[25] Nature. Availability of data, materials, code and protocols[EB/OL]. [2018 - 09 - 18]. https://www.nature.com/authors/policies/availability.html#data.
[26] Geoscience data journal. Author guidelines[EB/OL]. [2018 - 09 - 18]. https://rmets.onlinelibrary.wiley.com/hub/journal/20496060/about/author-guidelines.
[27] Scientific data. Data policies[EB/OL]. [2018 - 09 - 18]. https://www.nature.com/sdata/policies/data-policies.
[28] Elsevier. Research data[EB/OL]. [2018 - 09 - 18]. https://www.elsevier.com/about/policies/research-data.
[29] 袁洪林. 国际地学 SCI 期刊《Geochemistry, Geophysics, Geosystems》[J]. 岩矿测试,2012,31(5):928.
[30] Springer nature. Research data policy types[EB/OL]. [2018 - 09 - 19]. https://www.springernature.com/gp/authors/research-data-policy/data-policy-types/12327096.
[31] Scientific data. For authors[EB/OL]. [2018 - 09 - 19]. https://www.nature.com/sdata/publish/for-authors#format.
[32] Molecular ecology. Author guidelines[EB/OL]. [2018 - 09 - 19]. https://onlinelibrary.wiley.com/page/journal/1365294x/homepage/forauthors.html.
[33] Nature. Data availability statements and data citations policy: guidance for authors[EB/OL]. [2018 - 09 - 19]. https://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf.
[34] Figshare. Key benefits[EB/OL]. [2018 - 09 - 20]. https://figshare.com/services/publishers.
[35] Dryad. The repository: key features[EB/OL]. [2018 - 09 - 20]. https://datadryad.org/pages/repository.
[36] Mendeley Data. Discover mendeley data[EB/OL]. [2018 - 09 - 20]. https://data.mendeley.com/.
[37] PDB. Features & highlights[EB/OL]. [2018 - 09 - 20]. https://www.rcsb.org/#Subcategory-deposit_deposit-data.
[38] ScienceDaily. GenBank[EB/OL]. [2018 - 09 - 20]. https://www.sciencedaily.com/terms/genbank.htm.
[39] Wiley. Wiley's best practice guidelines on publishing ethics copyright and intellectual property[EB/OL]. [2018 - 09 - 21]. https://authorservices.wiley.com/ethics-guidelines/copyright-and-intellectual-property.html.
[40] EarthChem. EarthChem overview[EB/OL]. [2018 - 09 - 25]. http://www.earthchem.org/overview.
[41] Harvard dataverse. Subject: Medicine, Health and Life Sciences [EB/OL]. [2018 - 09 - 25]. https://dataverse.harvard.edu/dataverse/harvard? q = &fq0 = subject_ss%3A"Medicine%2C + Health + and + Life + Sciences" &types = dataverses%3Adatasets&sort = dateSort&order = desc.

作者贡献说明:

江洪:论文选题与设计,提出论文修改建议;
刘敬仪:论文框架构思,论文撰写与修改。

Foreign Journal Scientific Data Management Research

Jiang Hong[1] Liu Jingyi[1,2]
[1] Wuhan Library, Chinese Academy of Sciences, Wuhan 430071
[2] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190

Abstract: [**Purpose/significance**] It provides the inspiration and reference for domestic journals to carry out scientific data management, and puts forward journal scientific data management framework as well as reasonable development suggestions based on the domestic situation. [**Method/process**] This article selects 15 foreign journals that are good at scientific data management as research objects, and analyzes their characteristics in four aspects, including scientific data management related policies, scientific data submission, repository platform and open access. [**Result/conclusion**] The enlightenments are as follows: to formulate a complete scientific data management policy; establishing a detailed scientific data submission process; relating professional scientific data repository platform; promoting open access to scientific data actively.

Keywords: journal scientific data scientific data management framework